

CLAIMS

What is claimed is:

1. A method operative to automatically generate an index for a document, the
5 method comprising:
determining a sub-section delimiter definition;
searching the document to find occurrences of the defined sub-section
delimiter; and,
creating the index for the document from the found sub-section delimiter
10 occurrences.
2. The method operative to automatically generate an index for a document
of claim 1 wherein determining a sub-section delimiter comprises indicating at least
one of a font size, a font, a text string, a text location, a symbol, and a specific point
15 within the document.
3. The method operative to automatically generate an index for a document
of claim 1 wherein determining a sub-section delimiter comprises using a symbol
representing a demarcation point on a printed version of the document as the sub-section
20 delimiter.
4. The method operative to automatically generate an index for a document
of claim 1 wherein searching the document comprises:
generating an electronic version of the document; and,
25 searching the electronic version of the document for one of characters and
objects that match the defined sub-section delimiter .
5. The method operative to automatically generate an index for a document
of claim 4 wherein generating an electronic version of the document comprises:
30 scanning a printed version of the document to generate scan data, and,

performing one of optical character recognition functions and document recognition functions on the scan data to generate an electronic version of the document.

6. The method operative to automatically generate an index for a document of claim 1 further comprising:

displaying the created index;
checking that the displayed index is correct; and,
correcting the index.

7. The method operative to automatically generate an index for a document of claim 1 wherein determining a sub-section delimiter definition comprises:

selecting an exemplary sub-section title;
performing one of: document recognition and optical character recognition on the selected exemplary sub-section title, and
using at least one recognized property of the exemplary sub-section title as a subsection delimiter definition.

8. The method operative to automatically generate an index for a document of claim 1 wherein determining a sub-section delimiter definition comprises:

displaying a plurality of document pages on a user interface;
selecting at least one demarcation point on at least one of the plurality of pages ;
and,
using the at least one demarcation point as the defined sub-section delimiter.

9. A document processor operative to automatically generate an index for a document, the document processor comprising:

a document input device operative to provide an electronic version of a document;
a document storage device operative to store the electronic version of the document;

a delimiter searcher operative to search for and record information regarding occurrences of a defined delimiter within the electronic version of the document; and

a document divider operative to divide the document into sub-sections based on the recorded information regarding occurrences of the defined delimiter.

10. The document processor of claim 9 further comprising:

a user interface operative to transfer information between a document processor operator and portions of the document processor; and;

a delimiter designator module operative to communicate with the document processor operator through the user interface in order to generate at least one delimiter designation.

11. The document processor of claim 10 wherein the delimiter designator is operative to accept an indication of at least one of a font size, a font, a text string, a text location, a symbol, and a specific point within the document as a delimiter designation.

12. The document processor of claim 10 wherein the delimiter designator is operative to display a plurality of document portions on the user interface for the document operator to view while determining the at least one delimiter designation.

13. The document processor of claim 12 wherein the user interface is operative to receive demarcation point designations from the document processor operator and deliver the demarcation point designations to the delimiter designator as delimiter designations.

14. The document processor of claim 9 wherein the delimiter searcher is operative to search for a defined delimiter comprising a symbol selected from a barcode and a data glyph.

15. The document processor of claim 9 further comprising a print engine operative to print sub-sections of the document.

16. The document processor of claim 15, operating in a xerographic environment, wherein the print engine comprises a xerographic printer.

17. The document processor of claim 15 wherein the print engine comprises an inkjet printer.

18. A method for dividing a document into separate sections, the method comprising:

scanning the document to generate scanned document data;

performing recognition functions on the scanned document data to generate a recognized version of the document ;

defining a sub-section delimiter;

searching the recognized version to find occurrences of the defined sub-section delimiter; and,

using found sub-section delimiter occurrences to separate the document into the separate sections.

19. The method for dividing a scanned document into separate sections of claim 18 wherein defining a sub-section delimiter comprises at least one of building a sub-section delimiter from a list of predetermined potential sub-section delimiter components, performing statistical analysis on recognized characters to select characteristics that are most likely to be associated with sub-section delimiters, entering a sub-section delimiter through keyboard keystrokes, entering a sub-section delimiter by selecting symbols on a displayed portion of the electronic version of the document, and designating at least one demarcation point on at least one displayed portion of the electronic document to create a list of demarcation points to be used as a set of delimiter definitions.

20. The method for dividing a scanned document into separate sections of claim 18 wherein defining a sub-section delimiter comprises marking a paper version of the document with at least one special demarcation symbol prior to scanning the document.

5